

B. Tedeschi

1642

PH#71

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/402,713A

DATE: 01/11/2001
TIME: 13:47:34

Input Set : A:\seqlist-1619 0020001.txt
Output Set: N:\CRF3\01112001\I402713A.raw

ENTERED

RECEIVED

JAN 25 2001

TECH CENTER 1600/2000

4 <110> APPLICANT: Bussemakers, Marion J.G.
6 <120> TITLE OF INVENTION: PCA3, PCA3 Genes, and Methods of Use
8 <130> FILE REFERENCE: 1619.0020001
10 <140> CURRENT APPLICATION NUMBER: US 09/402,713A
11 <141> CURRENT FILING DATE: 2000-06-13
13 <150> PRIOR APPLICATION NUMBER: PCT/CA98/00346
14 <151> PRIOR FILING DATE: 1998-04-09
16 <150> PRIOR APPLICATION NUMBER: US 60/041,836
17 <151> PRIOR FILING DATE: 1997-04-10
19 <160> NUMBER OF SEQ ID NOS: 8
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 2037
25 <212> TYPE: DNA
26 <213> ORGANISM: Unknown Organism
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (379)..(531)
32 <220> FEATURE:
33 <223> OTHER INFORMATION: Description of Unknown Organism: PCA3
35 <220> FEATURE:
36 <221> NAME/KEY: unsure
37 <222> LOCATION: (1472)
38 <223> OTHER INFORMATION: May be any nucleotide.
40 <220> FEATURE:
41 <221> NAME/KEY: unsure
42 <222> LOCATION: (1517)
43 <223> OTHER INFORMATION: May be any nucleotide.
45 <220> FEATURE:
46 <221> NAME/KEY: unsure
47 <222> LOCATION: (1563)
48 <223> OTHER INFORMATION: May be any nucleotide.
51 <400> SEQUENCE: 1
52 agaagctggc atcagaaaaa cagaggggag atttgtgtgg ctgcagccga gggagaccag 60
54 gaagatctgc atgggtggaa ggaacctgat atacagagga attacaacac atatacttag 120
56 tgtttcaatg aacaccaaga taataagtg aagagctagt ccgctgtgag tctctcagt 180
58 gacacagggc tggatcacca tcgacggcac ttcttgagta ctacgtgcag caaagaaaga 240
60 ctacagacat ctcaatggca ggggtgagaa ataagaaagg ctgctgactt taccatctga 300
62 ggccacacat ctgctgaaat ggagataatt aacatcacta gaaacagcaa gatgacaata 360
64 taatglttaa gtatgtgac atg ttt ttg cac att tcc agc ccc ttt aaa tat 411
65 Met Phe Leu His Ile Ser Ser Pro Phe Lys Tyr
66 1 5 10
69 cca cac aca cag gaa gca caa aag gaa gca cag aga tcc ctg gga gaa 459
70 Pro His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg Ser Leu Gly Glu
71 15 20 25
73 atg ccc ggc cgc cat ctt ggg tca tcg atg agc ctc gcc ctg tgc ctg 507
74 Met Pro Gly Arg His Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu

RAW SEQUENCE LISTING DATE: 01/11/2001
 PATENT APPLICATION: US/09/402,713A TIME: 13:47:34

Input Set : A:\seqlist-1619 0020001.txt
 Output Set: N:\CRF3\01112001\I402713A.raw

75 30 35 40
 77 gtc cgg ctt gtc agg gaa gga cat tagaaaaatga attgatgtgt tccctaaagg 561
 78 Val Pro Leu Val Arg Glu Gly His
 79 45 50
 81 atggggcagga aaacagatcc tgttgtggat atttatttga acgggattac agatttgaaa 621
 83 tgaugtcaca aagttagcat taccatagag aggaaaacag acgagaaaat cttgatggct 681
 85 tcacaagaca tgcacacaac aaatgggaat actgtgatga catgaggcag ccaagctggg 741
 87 gaggagataa ccacggggca gagggtcagg attctggccc tgcgtccctaa actgtgcgtt 801
 89 cataaccaaa tcatttcata ttcttaaccc tcaaaacaaa gctgttgtaa tatctgatct 861
 91 ctacgggttc ttctggggcc aacatttctc atatatccag ccacactcat ttttaatat 921
 93 tagttcccg atctgtactg tgacctttct acactgtaga ataacattac tcattttgtt 981
 95 caaagaccct tctgttgtct gcttaatatg tagctgactg tttttccctaa ggagtgttct 1041
 97 gggccagggg atctgtgaac aggtggggaa gcactccaag atctttccag ggttataact 1101
 99 actagcacac agcatgatca ttacggagtg aattatctaa tcaacatcat cctcagtgtc 1161
 101 ttgtcccata ctgaaattca ttctccactt ttgtgcccct tctcaagacc tcaaatgtc 1221
 103 attccattaa taccacagga ttaacttttt tttttaacct ggaagaattc aatgttacat 1281
 105 gcagctatgg gaatttaatt acatattttg ttttccagt caaagatgac taagtctttt 1341
 107 atccctcccc ttgtttgat ttttttcca gtataaagt aaaatgctta gcttgtact 1401
 109 gaggctgtat acggcacagc ctctcccat cctccagcc ttatctgtca tcaccatcaa 1461
 111 ccccctccat nysacctaaa caaaatctaa cttgtaattc cttgaacatg tcaggncata 1521
 113 catttttctt tctgctgag aagctcttcc ttgtctctta antctagaat gatgtaaagt 1581
 115 ttggaataag ttgactatct tacttcatgc aaagaaggga cacatatgag attcatcatc 1641
 117 acatgagaca gcaaatacta aaagtgtaat ttgattataa gagtttagat aatatatga 1701
 119 aatgcaagak ccacagaggg aatgtttatg gggcacgttt gtaagcctgg gatgtgaagm 1761
 121 aaaggcaggg aacctcatag tatcttatat aatatacttc atttctctat ctctatcaca 1821
 123 atatccaaca agcttttccac agaattcatg cagtgcgaat ccccaaaggt aacctttatc 1881
 125 catttcatgg ttagtgctct ttagaatttt ggcaaatcat actggtcact tatctcaact 1941
 127 ttgagatgtg ttgtctcct tagttaattg aaagaaatag ggcactcttg tgagccactt 2001
 129 tagggttcac tcttgccaat aaagaattta caaaga 2037
 132 <210> SEQ ID NO: 2
 133 <211> LENGTH: 51
 134 <212> TYPE: PRT
 135 <213> ORGANISM: Unknown Organism
 137 <220> FEATURE:
 138 <223> OTHER INFORMATION: Description of Unknown Organism: PCA3
 140 <400> SEQUENCE: 2
 141 Met Phe Leu His Ile Ser Ser Pro Phe Lys Tyr Pro His Thr Gln Glu
 142 1 5 10 15
 144 Ala Gln Lys Glu Ala Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His
 145 20 25 30
 147 Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg
 148 35 40 45
 150 Glu Gly His
 151 50
 155 <210> SEQ ID NO: 3
 156 <211> LENGTH: 1872
 157 <212> TYPE: DNA
 158 <213> ORGANISM: Unknown Organism
 160 <220> FEATURE:

RAW SEQUENCE LISTING DATE: 01/11/2001
 PATENT APPLICATION: US/09/402,713A TIME: 13:47:35

Input Set : A:\seqlist-1619 0020001.txt
 Output Set: N:\CRF3\01112001\I402713A.raw

161 <223> OTHER INFORMATION: Description of Unknown Organism: PCA3
 163 <220> FEATURE:
 164 <221> NAME/KEY: unsure
 165 <222> LOCATION: (1307)
 166 <223> OTHER INFORMATION: May be any nucleotide.
 168 <220> FEATURE:
 169 <221> NAME/KEY: unsure
 170 <222> LOCATION: (1352)
 171 <223> OTHER INFORMATION: May be any nucleotide.
 173 <220> FEATURE:
 174 <221> NAME/KEY: unsure
 175 <222> LOCATION: (1398)
 176 <223> OTHER INFORMATION: May be any nucleotide.
 179 <400> SEQUENCE: 3
 180 agaagctggc atcagaaaaa cagaggggag atttgtgttg ctgcagccga gggagaccag 60
 181 gaagatctgc atggtgggaa ggacctgatg atacagaggt gagaaataag aaaggctgct 120
 182 gactttacca tctgaggcca cacatctgct gaaatggaga taattaacat cactagaaac 180
 183 agcaagatga caatataatg tctaaagtat gacatgtttt tgcacatttc cagccccctt 240
 184 aaatatccac acacacagga agcacaaaag gaagcacaga gatccctggg agaaatgcc 300
 185 ggccgccatc ttgggtcatc gatgagcctc gccctgtgcc tgggtccgct tgtgagggaa 360
 186 ggacattaga aaatgaattg atgtgttcct taaaggatgg gcaggaaaac agatcctggt 420
 187 gtggatatct atttgaacgg gattacagat ttgaaatgaa gtcacaaagt gagcattacc 480
 188 aatgagagga aaacagacga gaaaatcttg atggtctcac aagacatgca acaaacaaaa 540
 189 tggaaactct tgatgacatg aggcagccaa gctggggagg agataaccac ggggcagagg 600
 190 gtcaggattc tggccctgct gcttaactg tgcgttcata accaaatcat ttcatttttc 660
 191 taacctccaa aacaaagctg ttgtaatatc tgatctctac ggttccttct gggcccaaca 720
 192 ttctccatat atccagccac actcattttt aatatattag tcccagatct gtactgtgac 780
 193 ctttctacac ttagaataaa cattactcat tttgttcaaa gaccttcgtt gttgctgct 840
 194 aatatgtagc tgactgtttt tcttaaggag tgttctggcc caggggatct gtgaacaggc 900
 195 tgggaagcat ctcaagatct ttccagggtt atacttacta gcacacagca tgatcattac 960
 196 ggagtgaatt atctaataca catcatcttc agtgtctttg cccatactga aattcatttc 1020
 197 ccacttttgt gccattctc aagacctcaa aatgtcatc cattaatatc acaggattaa 1080
 198 cttttttttt taacctggaa gaattcaatg ttacatgcag ctatgggaat ttaattacat 1140
 199 attttgtttt ccagtgcaaa gatgactaag tcttttatcc ctccccttg tttgattttt 1200
 200 ttccagtat aaagttaaaa tgccttagct tgtactgagg ctgtatacag cacagcctct 1260
 201 ccccatccct ccagccttat ctgtcatcac catcaacccc tcccatnysa cctaaacaaa 1320
 202 atctaacttg taattccttg aacatgtcag gncatacatt rtctctctg cctgagaagc 1380
 203 tcttctctgt ctcttaantc tagaatgatg taaagttttg aataagttga ctatcttact 1440
 204 tcatgcaaaag angggacaca tatgagatc atcatcacat gagacagcaa atctctaaaag 1500
 205 tgtaatttga ttataagagt ttatgataat atatgaaatg caagakccac agagggaaatg 1560
 206 ttatgggggc acgtttgtta gccctgggatg tgaagmaaag gcagggaacc tcatagtatc 1620
 207 ttatataata tacttcattt ctctatctct atcacaatat ccaacaagct ttccacagaa 1680
 208 ttcatgcagt gcaaatcccc aaaggtaacc ttatccatt tcatggtgag tgcgctttag 1740
 209 aattt.tggca aalcalaactg gtcacttalc tcaactttga gatgtglttg tcttctgtagt 1800
 210 taattgaaaag aaatagggca ctcttctgag ccacttttag gttcactcct ggcaataaag 1860
 211 aatttcaaaa ga 1872
 214 <210> SEQ ID NO: 4
 215 <211> LENGTH: 820
 216 <212> TYPE: DNA

RAW SEQUENCE LISTING DATE: 01/11/2001
 PATENT APPLICATION: US/09/402,713A TIME: 13:47:35

Input Set : A:\seqlist-1619 0020001.txt
 Output Set: N:\CRF3\01112001\I402713A.raw

```

217 <213> ORGANISM: Unknown Organism
219 <220> FEATURE:
220 <223> OTHER INFORMATION: Description of Unknown Organism: PCA3
222 <400> SEQUENCE: 4
223 agaaagctggc atcagaaaaa cagaggggag atttgtgttg ctgcagccga gggagaccag 60
224 gaagatctgc atggtgggaa ggacctgatg atacagaggt gagaaalaag aaaggctgct 120
225 gactttacca tctgagggca cacatctgct gaaatggaga taattacat cactagaaac 180
226 agcaagatga caatataatg tctaagtaqt gacatgtttt tgcacatttc cagccctttt 240
227 aaatatccac acacacagga agcacaaaag gaagcacaga gatccctggg agaaatgccc 300
228 ggcggccatc ttgggtcatc gatgagcctc gccctgtgac tgggtccgct tgtgagggaa 360
229 ggacattaga aatgaattg atgtgttctt taaaggatgg gcaggaaaac agatcctggt 420
230 gtggatattt atttgaacgg gattacagat ttgaaatgaa gtcacaaagt ggcattacc 480
231 aatgagagga aaacagacga gaaaatcttg atggtttcac aagacatgca acaaacaaaa 540
232 tggaaatactg tgatgacatg aggcagccaa gctggggagg agataaccac ggggcagagg 600
233 gtcagygatc tggcctctgt gccataactg tgcgttcata accaaatcat ttcatatttc 660
234 taacctcaca aacaaagctg ttgtaatac tgatctctac ggttccttct gggcccaaca 720
235 ttctccatat atccagccac actcattttt aatatttagt tcccagatct gtactgtgac 780
236 cttttctcac ttagaataaa cattactcat ttgtttcaaa 820
239 <210> SEQ ID NO: 5
240 <211> LENGTH: 11
241 <212> TYPE: PRT
242 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: Description of Artificial Sequence: antigenic
246 peptide
248 <400> SEQUENCE: 5
249 His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg
250 1 5 10
253 <210> SEQ ID NO: 6
254 <211> LENGTH: 3582
255 <212> TYPE: DNA
256 <213> ORGANISM: Unknown Organism
258 <220> FEATURE:
259 <221> NAME/KEY: CDS
260 <222> LOCATION: (401)..(553)
262 <220> FEATURE:
263 <223> OTHER INFORMATION: Description of Unknown Organism: PCA3
265 <400> SEQUENCE: 6
266 acagaagaaa tagcaagtgc cgagaagctg gcacacagaa aacagagggg agatttgtgt 60
267 ggtgcagcc gagggagacc aggaagatct gcatgggtgg aaggacctga tgatacagag 120
270 gaattacaac acatatactt agtgtttcaa tgaacaccaa gataaataag tgaagagcta 180
272 gtccgctgtg agtctcctca gtgacacagg gctggatcac catcgacggc actttctgag 240
274 tactcagtg agcaaaagaa gactacagac atctcaatgg caggggtgag aaataagaaa 300
276 ggctgctgac ttaccatct gaggccacac atctgctgaa atygagataa ttaacalcac 360
278 tagaaacagc aagatgacaa tataatgtct aagtagtgac atg ttt ttg cac att 415
279 Met Phe Leu His Ile
280 1 5
282 tcc agc ccc ttt aaa tat cca cac aca cag gaa gca caa aag gaa gca 463
283 Ser Ser Pro Phe Lys Tyr Pro His Thr Gln Glu Ala Gln Lys Glu Ala

```

RAW SEQUENCE LISTING

DATE: 01/11/2001

PATENT APPLICATION: US/09/402,713A

TIME: 13:47:35

Input Set : A:\seqlist-1619 0020001.txt

Output Set: N:\CRF3\01112001\I402713A.raw

```

284          10          15          20
286  cag aga tcc ctg gga gaa atg ccc gcc cgc cat ctt ggg tca tgg atg 511
287  Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His Leu Gly Ser Ser Met
288          25          30          35
290  agc ctc gcc ctg tgc ctg gtc cgg ctt gtg agg gaa gga cat 553
291  Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg Glu Gly His
292          40          45          50
294  tagaaaatga attgatgtgt tccctaaagg atgggcaggga aaacagatcc tgttgtggat 613
296  atttatattga acygggttac agatttgaaa tgaagtccaca aagtgagcat taccaatgag 673
298  aggaaaacag acgagaaaat cttgatggct tcacaagaca tgcacaaaac aaagtggat 733
300  actgtgatga catgaggcag ccaagctggg gaggagataa ccacggggca gagggtcagg 793
302  attctyggcc tgetgcctaa actgtgcgtt cataaccaaa tcatttcaata ttcttaaccc 853
304  tcaaaacaaa gctgttgtta tatctgatct ctacgyttcc ttctggqccc aacattctcc 913
306  atatatccag ccacactcat ttttaatat tagttccag atctgtactg tgacctttct 973
308  acactytaga ataacattac tcattttgtt caaagacctt tegtgttgtt gcttaatatg 1033
310  tagctgactg tttttcctaa ggaagtgtct ggcacagggg atctgtgaac aggtgagaa 1093
312  gcattcctag atctttccag ggttatactt actagcacac agcatgatca ttacggagtg 1153
314  aattatctaa tcaacatcat cctcagtgtc tttgcccata ctgaaattca ttcccaactt 1213
316  ttgtgcccct tctcaagacc tcaaaatgtc attccattaa tatcacagga ttaacttttt 1273
318  tttttaacct ggaagaattc aatgtttacat qcagctatgg gaatttaatt acatattltg 1333
320  ttttccagtg caaagatgac taagtccctt atccctcccc ttgtttgat ttttttcca 1393
322  gtataaagtt aaaatgctta gccttgtact gaggtgtat acagcacagc ctctcccat 1453
324  cccctcagcc ttatctgtca tcaaccataa cccctcccat accacctaaa caaaatctaa 1513
326  ctgttaattc ctgtgaacatg tcaggacata cattattcct tctgcctgag aagctcttcc 1573
328  ttgtctctta aatctagaat gatgtaaagt tttgaataag ttgactatct tacttcatgc 1633
330  aaagaaggga cacatagag attcatcatc acatgagaca gcaaatacta aaagtgtaat 1693
332  ttgatataaa gagtttagat aaatatatga aatgcaagag ccacagaggg aatgtttatg 1753
334  gggcacgttt gtaagcctgg gatgtgaagc aaaggcaggg aacctcatag tatcttatat 1813
336  aatatacttc atttctctat ctctatcaca atatccaaac agcttttcc agaatctatg 1873
338  cagtgcaaat ccccaaaagt aacctttatc catttcatgg tgaagtgcgt ttagaatttt 1933
340  ggcaaatcat actgttccat tatctcaact ttgagatgtg ttgtctcttg tagttaattg 1993
342  aaagaaatag ggcactcttg tgagccactt tagggttcac tccgtggcaat aaagaattta 2053
344  caaagagcta ctcaggacca gttgttaaga gctctgtgtg tgtgtgtgtg tgtgtgtgag 2113
346  tgtacatgcc aaagtgtgoc tctctctctt gacccattat ttccagactta aaacaagcat 2173
348  gttttcaaat ggcactatga gctgccaatg atgtatcacc accatattct attattctcc 2233
350  agtaaatgtg ataataatgt catctgttaa cataaaaaaa gtttgacttc acaaaagcag 2293
352  ctggaaatgg acaaccacaa tatgcataaa tctaaactct accatcagct acacactgct 2353
354  tgacatatat tgttagaagc acctcgcatl tgtgggttct cttaagcaaa atacttgcac 2413
356  taggtctcag ctggggctgt gcctcaggcg gtttgagaaa tattcaattc tcagcagaag 2473
358  ccagaatttg aatlcctca tcttttagga atcatttacc aggtttlgag aggattcaga 2533
360  cagctcaggt gctttcacta atgtctctga actctgttcc ctcttttgtt tcatggatag 2593
362  tccaataaat aatgtttatct ttgaactgat gctcatagga gagaatataa gaactctgag 2653
364  tgatatcaac attagggatt caaagaaata tttagattta gctcacactg gtcanaagga 2713
366  accaagatac aaagaactct gagctgtcat cgtcccccac tctgtgagcc acaaccaaca 2773
368  gcaggaccoca acgcatgtct gagatcctta aatcaaggaa accagtgta tgagtgaat 2833
370  tctctatata tggatgctag cttctyggca tctctggctc tccctttgac acatattage 2893
372  ttctagcctt tgcctccag acttttatct ttctccaac acatcgctta ccaatcctct 2953
374  ctctgctctg ttgclttgga ctctcccaac agaatttcaa cgaactctca gtcctttctt 3013
376  ccactccccc cactaacctg aattgcctag acccltattt ttattatttt ccaatagatg 3073

```

VERIFICATION SUMMARY DATE: 01/11/2001
PATENT APPLICATION: US/09/402,713A TIME: 13:47:36

Input Set : A:\seqlist-1619 0020001.txt
Output Set: N:\CRF3\01112001\I402713A.raw

L:111 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:113 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:201 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:202 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:203 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3